NPS-53-84-0006

NAVAL POSTGRADUATE SCHOOL

Monterey, California

Evaluating A^T*D*A for Sparse Matrices: Analysis

by

Amnon Gonen

June 1984

Technical Report for Period

May 1983 - June 1984

Approved for public release: distribution unlimited

Prepared for: Chief of Naval Research
Arlington, VA 22217

NAVAL POSTGRADUATE SCHOOL
Monterey, California 93943

R. H. SHUMAKER
Commodore, U. S. Navy
Superintendent

D. A. SCHRADY
Provost

This report was prepared by:

AMNON GONEN
Adjunct Professor of
Mathematics

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NPS-53-84-0006	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Evaluating A^T*D*A for Sparse Matrices: Analysis		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Amnon Gonen		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93943		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61152N: RR000-01-10 N0001484WR41001
11. CONTROLLING OFFICE NAME AND ADDRESS NPS Foundation Research Program NPS, Monterey, CA 93943		12. REPORT DATE June 1984
		13. NUMBER OF PAGES 23
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Chief of Naval Research Arlington, VA 22217		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Sparce Matrix, Hessian evaluation, Optimization

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The evaluation of the matrix product A^T*A or A^T*D*A, where A is an mxn real matrix and D an mxm diagonal matrix, is a fundamental operation for many algorithms. We analyze the evaluation of A^T*A for several configurations of sparse matrices A all of which have the same sparsity. The complexity of the evaluation is estimated, and application to certain problems of optimization are given.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

S/N 0102- LF- 014- 6601

Evaluating $A^T \cdot D \cdot A$ for Sparse Matrices: Analysis

Amnon Gonen
Naval Postgraduate School

ABSTRACT

The evaluation of the matrix product $A^T \cdot A$ or $A^T \cdot D \cdot A$,where A is an $m \times n$ real matrix and D an $m \times m$ diagonal matrix, is a fundamental operation for many algorithms. We analyze the evaluation of $A^T \cdot A$ for several configurations of sparse matrices A all of which have the same sparsity. The complexity of the evaluation is estimated, and application to certain problems of optimization are given.

Key Words: Sparce Matrix, Hessian evaluation, Optimization

1. INTRODUCTION

Many fundamental algorithms in numerical analysis include the evaluation of $A^T \cdot A$ or $A^T \cdot D \cdot A$, where A is a real $m \times n$ matrix and D is a diagonal $m \times m$ matrix. Examples are given in papers on factorization of matrices or problems of minimization in which the Hessian has this form (Gay [1] Gonen & Avriel [3]). The extended use of this product motivates the question of reducing its complexity.

This research was partially supported by the NPS Foundation Research Program.

The purposes of this paper are:

1. To relate the computational complexity of $A^T \cdot D \cdot A$ to the sparsity rate of the matrix A .

2. For a given sparsity rate, to distinguish between the worst and the best case.

3. To provide an application of these results.

The problem of multiplying a transpose of a sparse matrix by itself was discussed in several books and papers e.g. George & Liu [2] in which they include the number of operations required for this multiplication. Gustavson [4] proposed an optimal algorithm for multiplying two sparse matrices $A \cdot B$ where $A \in R^{n \times m}$ and $B \in R^{m \times k}$, proving that the number of multiplication N satisfies $0 \leq N \leq nmk$. However, the connection between the number of operation and the sparsity rate of the matrices was not discussed.

Apparently, it seems that this question has only theoretical meaning since the matrix A is provided and therefore the number of operations is known. However, in this paper we will see there exist some cases in which the configuration of this matrix A can be designed by the user. In these cases it make sense to analyze this product in order to reduce the number of operations.

In section 2 of this paper, we present the computational complexity of $A^T \cdot D \cdot A$ for several sparsity patterns of A. In this section, we establish our results on the assumption that the number of nonzero elements of

the matrix A is provided. We demonstrate the best and the worst case, showing that in the best case, the nonzero elements are divided homogeneously among the rows of A, while in the worst case, these nonzero elements are confined in a limited number of rows.

In section 3 we provide an example from optimization theory, in which the matrix A is dense and by applying the results of section 2 we minimize the number of multiplication in the evaluation of the Hessian.

In this paper, all vector spaces are finite dimensional and vectors are column vectors. The space of all $n \times m$ matrices is denoted by $R^{n\times m}$; the nonnegative orthant of the Euclidean space R^n is denoted by R^n_+ ; the subset of all integer vectors in R^n is denoted by I^n, and its nonnegative orthant by I^n_+. For a matrix A we denote by $a_{i\bullet}$ and $a_{\bullet j}$ the i-th row and the j-th column respectively. The transpose of A is denoted by A^T. By the norm $||x||$ we mean the Euclidean norm. For a real number r its integer part is denoted by $\lfloor r \rfloor$. Finally, the number of elements in the set B is denoted by $|B|$, and the number of zero elements in a matrix A is denoted by $Z(A)$.

2. THE COMPUTATIONAL COMPLEXITY OF $A^T \cdot D \cdot A$.

Let A be in $R^{m\times n}$ with N nonzero elements. The ratio $\frac{N}{mn}$ is called <u>the sparsity rate</u> of the matrix A and denoted by $\sigma(A)$. In this section we assume that the sparsity rate of the matrix $A \in R^{m\times n}$ is provided and that

each row of A includes, at least, one nonzero element. We concentrate on the sparsity pattern of A, looking for the best and the worst cases, by means of the number of operations required to compute $A^T \cdot D \cdot A$ where D is a diagonal matrix $D \in R^{m \times m}$. We begin our exploration in the worst case, in which the configuration of A implies the maximum number of multiplication. Let us denote by m_i the number of nonzero elements in a_{i*} , thus

$$\sum_{i=1}^{m} m_i = N. \tag{2.1}$$

Our first Lemma provides us the number of operations (multiplications) required to accomplish the product $A^T \cdot A$.

Lemma 2.1: Let $A \in R^{m \times n}$ be a given sparse matrix, then the product $A^T \cdot A$ can be computed using

$$\tfrac{1}{2} \sum_{i=1}^{m} m_i \cdot (m_i + 1) \tag{2.2}$$

multiplications.

Proof: The product $A^T \cdot A$ can be rewritten as a sum of m rank 1 matrices

$$A^T \cdot A = \sum_{i=1}^{m} a_{i*} \, a_{i*}^T \tag{2.3}$$

The rank 1 matrices $a_{i*} \cdot a_{i*}^T$ are symmetric. Each nonzero element $a_{i,j}$ of the vector a_{i*} is multiplied by all other nonzero elements $a_{i,k}$ for $k \geq j$. Therefore, the number of multiplication is

$$\sum_{i=1}^{m_i} i = \tfrac{1}{2} m_i \cdot (m_i + 1) \tag{2.4}$$

combining (2.3) with (2.4) yields the proof of the lemma.

▪

From the proof above, it can easily be seen that the number of additions are approximately the same as the number of multiplications since each term $a_{i,k} \cdot a_{k,j}$ is accumulated into the result matrix C; $C = A^T \cdot A$.

Corollary 2.1: Let $A \in R^{m \times n}$ be a sparse matrix and $D \in R^{m \times m}$ a diagonal matrix then the product $A^T \cdot D \cdot A$ can be computed by

$$\tfrac{1}{2} \sum_{i=1}^{m} m_i \cdot (m_i + 1) + N \tag{2.5}$$

multiplications.

Proof: We first compute $\bar{A} = D \cdot A$ which requires N multiplications and then substituting $\bar{a}_{i\cdot}^T$ by $a_{i\cdot}^T$ in (2.3) yields the proof of the corollary.

In order to find the sparsity pattern which yields the worst case, we have to maximize (2.2) provided (2.1) and all m_i are positive integers. Since the difference between (2.2) and (2.5) is N, it is enough to explore the worst case for the product $A^T \cdot A$ that will yield the same result for $A^T \cdot D \cdot A$. Consequently, a new problem can be formulated as follow.

(A1) $$\max \sum_{i=1}^{m} \frac{m_i \cdot (m_i + 1)}{2} \tag{2.6}$$

subject to the constraint

$$\sum_{i=1}^{m} m_i = N$$

and

$$1 \le m_i \le n \; ; \; m_i \in I^1 \tag{2.7}$$

This problem can be reduced to maximizing $\sum_{i=1}^{m} m_i^2$ under the same constraints. Defining $x_i = m_i - 1$ yields the following problem

(A2) $$\max ||x||^2 \tag{2.8}$$

subject to the constraint

$$\sum_{i=1}^{m} x_i = N - m \tag{2.9}$$

and

$$0 \le x_i \le n - 1 \, , \; x_i \in I^1 \tag{2.10}$$

We will prove that since the objective function is convex its maximum is attained at a boundary point. An integer vector $x \in I^m$ is called a **boundary point** of problem (A2) if there exists a set $J = \{j_1, \ldots, j_\vartheta\} \subset L = \{1, \ldots, m\}$ and a unique $j_0 \in L - J$ such that

$$x_i = \begin{cases} n - 1 & i \in J \\ (N - m) - \vartheta(n - 1) & i = j_0 \\ 0 & otherwise \end{cases} \tag{2.11}$$

where

$$\vartheta = \left\lfloor \frac{N - m}{n - 1} \right\rfloor . \tag{2.12}$$

In this case the vector $\bar{x} = x + e$ where $e = (1, \ldots, 1)$ is a boundary point of problem (A1).

Fortunately, from the symmetric property of the objective function, the optimal value does not depend on the selected boundary point. Hence

$$\sum_{i=1}^{m} x_i^2 = \vartheta\cdot(n-1)^2 + [(N-m) - \vartheta\cdot(n-1)]^2. \tag{2.13}$$

To prove that (2.11) is the solution of problem (A2) we need the following lemma

Lemma 2.2 : Consider the integer problem

(A3) $$\max_{x \in I^n} ||x||^2 \tag{2.14}$$

subject to the constraints

$$\sum_{i=1}^{n} x_i = K \tag{2.15}$$

$$0 \le x_i \le M \tag{2.16}$$

where K and M are positive integers, $M \le K$. Problem (A3) has a solution x^* satisfying

$$||x^*||^2 = \vartheta M^2 + (K - \vartheta M)^2 \tag{2.17}$$

where ϑ is the integer part of $\frac{K}{M}$ denoted by $\left\lfloor \frac{K}{M} \right\rfloor$, if and only if

$$n \cdot M \ge K. \tag{2.18}$$

Moreover, if (2.18) holds then every solution of problem (A3) satisfies (2.17).

Proof: It is immediate that if $n \cdot M < K$ then there is no feasible solution to problem (A3). Therefore, let us assume that (2.18) holds and prove this lemma by induction on the dimension of x. If $n = 1$, then from (2.15) we have $x = K \leq M$. If $K < M$ then $\vartheta = 0$ and if $K = M$ then $\vartheta = 1$. In both cases (2.17) is satisfied. Assuming the assertion is true for all n, $n \leq m-1$. Let us denote

$$F(m,M,K) = \max\{\,||x||^2 : \sum_{i=1}^{m} x_i = K \,;\, 0 \leq x_i \leq M \,;\, x \in I^m\} \qquad (2.19)$$

hence

$$F(m,M,K) = \max_{1 \leq x_m \leq M} \{F(m-1,M,K-x_m) + x_m^2 : x_m \in I_+^1\}. \qquad (2.20)$$

Since by the induction assumption, (2.17) holds for $m-1$

$$F(m,M,K) = \max_{1 \leq x_m \leq M} \{\overline{\vartheta} M^2 + (K - x_m - \overline{\vartheta} M)^2 + x_m^2 \;:\; x_m \in I_+^1 \,;\, \overline{\vartheta} = \left\lfloor \frac{K - x_m}{M} \right\rfloor\}. \qquad (2.21)$$

Let $\rho = K - \vartheta M$ where $\vartheta = \left\lfloor \frac{K}{M} \right\rfloor$ then $0 \leq \rho \leq M$. (ρ is the remainder of dividing K by M) Consider the maximization problem (2.21) in two cases:

1. $0 \leq x_m \leq \rho$.

In this case $\overline{\vartheta} = \vartheta$ and the problem is

$$\max_{0 \leq x_m \leq \rho} \{\vartheta M^2 + (K - \vartheta M)^2 - 2(K - \vartheta M) \cdot x_m + 2x_m^2 : \; x_m \in I_+^1\}. \qquad (2.22)$$

Substituting $K - \vartheta M$ by ρ , yields the maximization of

$$\vartheta M^2 + \rho^2 - 2\rho x_m + 2x_m^2 \qquad (2.23)$$

subject to the constraint $0 \leq x_m \leq \rho$ and $x_m \in I_+^1$. The maximum of (2.23) is

attained at $x_m = \rho$ or $x_m = 0$, and

$$||x||^2 = \vartheta M^2 + \rho^2 = \vartheta M^2 + (K - \vartheta M)^2. \tag{2.24}$$

2. $\rho < x_m \leq M$.

In this case $\overline{\vartheta} = \vartheta - 1$ and the problem is

$$\max_{\rho \leq x_m \leq M} \{ (\vartheta - 1)M^2 + (\rho + M)^2 - 2(\rho + M)x_m + 2x_m^2 \mid x_m \in I_+^1 \} \tag{2.25}$$

using the same arguments as in the first case, the maximum is attained at $x_m = M$, thus

$$||x||^2 = (\vartheta - 1)M^2 + (\rho + M)^2 - 2\rho M = \vartheta M^2 + \rho^2 = \vartheta M^2 + (K - \vartheta M)^2. \tag{2.26}$$

In both cases (2.17) holds, which complete our proof.

▫

Applying Lemma 2.2 to problems (A1) and (A2) yields the following conclusion.

Corollary 2.3: Every $x \in I_+^n$ satisfying (2.11) is a solution to problem (A2).

Proof: Suppose $x \in I_+^n$ satisfies (2.11), which mean that (2.13) holds. Substituting $M = n - 1$ and $K = N - m$ in Lemma 2.2 implies that (2.17) and (2.13) are the same, and Lemma 2.2 implies that x is a solution of problem (A2).

▪

A solution to problem (A1) can be established by setting

$$m_i = \begin{cases} n & i \in J \\ N - m - \vartheta(n-1) + 1 & i = j_0 \\ 1 & otherwise \end{cases} \tag{2.27}$$

where ϑ satisfies (2.12) and J is a set of indices $|J| = \vartheta$ and j_0 is an index not in J. The computational complexity of the product $A^T \cdot A$ for the worst case is established by substituting (2.27) into (2.6)

$$\mu_{wc} = \tfrac{1}{2}[\vartheta n^2 + (N - m - \vartheta(n-1) + 1)^2 + m - \vartheta - 1 + N]. \tag{2.28}$$

It can be seen that in the worst case some of the m_i-th achieve the upper bound n , the others are zero and only one of the m_i-th is somewhere between 0 and n. This mean that the matrix A has as many full rows as possible, the rest of the rows have one element, and one row contains the remaining nonzero elements of N.

In the next Lemma a new bound for the computational complexity is presented which enable us to relate the sparsity rate and the mathematical effort.

Lemma 2.4 The computational complexity of the product $A^T \cdot A$ can be bounded by

$$\mu_{wc} \le \tfrac{1}{2}(n(N-m) + 2N). \tag{2.29}$$

Proof: Let us denote by

$$\varphi(k) = (k \cdot n^2 + [(N-m) - k \cdot (n-1) + 1]^2 + m - k - 1 + N). \tag{2.30}$$

The first assertion is that

$$\varphi(\left\lfloor\frac{N-m}{n-1}\right\rfloor)\le\varphi(\frac{N-m}{n-1}). \qquad (2.31)$$

If we denote by $\zeta=\frac{N-m}{n-1}-\left\lfloor\frac{N-m}{n-1}\right\rfloor$, then $0\le\zeta<1$. A straightforward calculation yields that

$$\varphi(\frac{N-m}{n-1})=(N-m)\cdot(n+1)+m+N. \qquad (2.32)$$

Hence

$$\varphi(\frac{N-m}{n-1})-\varphi(\frac{N-m}{n-1}-\zeta)=(N-m)\cdot(n+1)+m+N- \qquad (2.33)$$

$$-(\frac{N-m}{n-1}n^2-\zeta n^2+[N-m-\frac{N-m}{n-1}(n-1)+\zeta(n-1)+1]^2+m-\frac{N-m}{n-1}+\zeta-1+N)=$$

$$=(N-m)\cdot(n+1)+m+N-[\frac{N-m}{n-1}(n^2-1)+m+N-\zeta n^2+(\zeta(n-1)+1)^2+\zeta-1]=$$

$$=\quad \zeta n^2-(\zeta(n-1)+1)^2-\zeta+1=\zeta\cdot(1-\zeta)\cdot(n-1)^2$$

since $0\le\zeta<1$ the last expression is nonnegative which prove our first assertion. The rest of the proof is established by the following:

$$\mu_{uc}=\tfrac{1}{2}\varphi(\left\lfloor\frac{N-m}{n-1}\right\rfloor)\le\tfrac{1}{2}\varphi(\frac{N-m}{n-1})=\tfrac{1}{2}[n(N-m)+2N]. \qquad (2.34)$$

∎

As we can see, (2.29) provides us an elegant bound for the computational complexity of the worst case. This bound is a good approximation to the computational complexity when $\frac{N-m}{n-1}$ is close to its integer part. The difference between the mathematical effort of computing $A^T\cdot A$ in the worst case and this bound is actually provided in the right hand side of

(2.33) and it is

$$\tfrac{1}{2}\zeta\cdot(1-\zeta)\cdot(n-1)^2 \tag{2.35}$$

where ζ is the fraction part of $\dfrac{N-m}{n-1}$.

The bound in (2.29) can be expressed as a function of the sparsity rate by using the definition

$$\sigma(A) = \frac{N}{mn} \tag{2.36}$$

which leads to the following equality

$$\mu_{wc} \le \tfrac{1}{2}[n(N-m)+2N] = \tfrac{1}{2}nm(\sigma(A)(n+2)-1). \tag{2.37}$$

It is interesting to observe the connection between the bound in (2.29) and the mathematical effort to accomplish $A^T\cdot A$ without using sparsity method which is

$$\tfrac{1}{2}n\cdot(n+1)\cdot m. \tag{2.38}$$

The difference between (2.38) and (2.29) can be established by expanding these two formulas achieving

$$\tfrac{1}{2}n(n+1)m - \tfrac{1}{2}[(N-m)n+2N] = \tfrac{1}{2}(n+2)(m\cdot n - N). \tag{2.39}$$

Dividing and multiplying the right hand side of (2.39) by mn yield the following expression for the difference

$$\tfrac{1}{2}(n+2)\cdot m\cdot n\cdot(1-\sigma(A)) \tag{2.40}$$

where $\sigma(A)$ is the sparsity rate of A.

2.2 The best case

In our discussion, we call the case in which we need the minimum number of multiplication to produce $A^T \cdot A$ provided that there are N nonzero elements in A **the best case** . The number of operations in the best case can be derived by minimizing

$$\sum_{i=1}^{m} \frac{m_i \cdot (m_i + 1)}{2} \tag{2.41}$$

subject to (2.6) and (2.7). Without the integer restriction it is immediate that since the objective function is convex, the solution will be the arithmetic mean, that is, for all i , $m_i = \frac{N}{m}$. The restriction that all the m_i have to be integral yields the solution

$$m_i^* = \begin{cases} \left\lfloor \frac{N}{m} \right\rfloor & i \in J \\ \left\lfloor \frac{N}{m} \right\rfloor + 1 & i \in L - J \end{cases} \tag{2.42}$$

where $L = \{1,\ldots,m\}$, $J \subseteq L$ and $|L - J| = N - \left\lfloor \frac{N}{m} \right\rfloor \cdot m$. Consequently, the number of multiplication in the best case is

$$\mu_{bc} = \sum_{i=1}^{m} \frac{m_i^* \cdot (m_i^* + 1)}{2} = \tfrac{1}{2}(\left\lfloor \frac{N}{m} \right\rfloor + 1) \cdot (2N - m \left\lfloor \frac{N}{m} \right\rfloor). \tag{2.43}$$

In order to present the magnitude of the difference between the worst and the best case, let us assume that $\frac{N}{m}$ and $\frac{N-m}{n-1}$ are integers. In this case (2.29) holds with equality and

$$\mu_{bc} = \frac{1}{2}\,\frac{N}{m}(N + m). \tag{2.44}$$

Subtracting μ_{bc} from μ_{wc} yield

$$\mu_{wc} - \mu_{bc} = \frac{1}{2}(nN - nm + N - \frac{N^2}{m}) = \tag{2.45}$$

$$= \frac{1}{2}n(N - m)(1 - \sigma(A)).$$

If we take, for example, $N = \frac{1}{2}m(n+1)$ the difference will be $\frac{m(n-1)^2}{8}$ while $\mu_{bc} = \frac{m(n+1)(n+3)}{4}$. That means that μ_{wc}, for large n, is approximately 50% more than μ_{bc}.

3. APPLICATION

In this section we present an example in which the product $A^T \cdot D \cdot A$ is required where D is a diagonal matrix and the pattern of A can be designed in order to reduce the computational effort. Since we are discussing the number of zeroes in matrices, let us denote by $Z(A)$ the number of zero elements in the matrix A. Consider the problem introduced by Gay [1]

$$\text{(P1)} \qquad \min \varphi(x) = \sum_{i=1}^{m} \rho_i(r_i(x)) \tag{3.1}$$

where $r_i : R^n \to R$, $\rho_i : R \to R$ and $m \geq n$. Very often $r(x) = (r_1(x), \ldots, r_m(x))$ is a linear function of x , (see for example Gonen & Avriel [3], or the least square problem in Gay [1]) which mean

$$r(x) = A \cdot x - b. \tag{3.2}$$

In this case, the gradient and Hessian of φ have particularly simple forms

$$\nabla\varphi(x) = A^T \cdot \rho'(r(x)) \tag{3.3}$$

$$\nabla^2\varphi(x) = A^T \cdot D \cdot A \tag{3.4}$$

where

$$\rho'(r(x)) = [\rho'_1(r_1(x)), \ldots, \rho'_m(r_m(x))] \tag{3.5}$$

and

$$D = diag[\rho''_1(r_1(x)), \ldots, \rho''_m(r_m(x))] \tag{3.6}$$

is the diagonal matrix with diagonal elements $\rho''_i(r_i(x))$. Since we have a simple analytic presentation of the gradient and Hessian , it is reasonable

to consider using Newton method to construct a sequence of iterates which, under reasonable conditions, converge to a local minimizer. This mean that the product $A^T \cdot D \cdot A$ will be used each iteration and very often this computation is the most expensive part of the algorithm. The main idea is to accomplish an initial preparation step by factoring

$$A = B \cdot Q \tag{3.7}$$

where $Q \in R^{n \times n}$ is nonsingular and $B \in R^{m \times n}$ has $(n^2 - n)$ zeroes in it $(Z(B) = n)$. The next step is to substitute Qx by y in (3.7) leading to the problem

$$\text{(P2)} \qquad \min \varphi(x) = \sum_{i=1}^{m} \rho_i(r_i(x)) \tag{3.8}$$

where

$$r(y) = By - b. \tag{3.9}$$

To establish the connection between the two problems, let us introduce the following Lemma:

Lemma 3.1: A point x^* satisfies sufficient conditions for minimum of problem P1 with $r(x)$ defined by (3.2) if and only if $y^* = Qx$ satisfies sufficient conditions for minimum of problem P2.

Proof: The sufficient conditions for minimum of problem P1, where $r(x)$ satisfies (3.2), are:

$$A^T \cdot \nabla \varphi(Ax^*) = 0 \tag{3.10}$$

$$z^T \cdot A^T \cdot \nabla^2 \varphi(Ax^*) Az > 0 \tag{3.11}$$

for all $z \neq 0$. Since $A = B \cdot Q$ where Q is a nonsingular matrix (3.10) is equivalent to

$$B^T \nabla \varphi(By^*) = 0 \tag{3.12}$$

and (3.11) can be rewritten as

$$z^T \cdot Q^T \cdot B^T \cdot \nabla^2 \varphi(By^*) B \cdot Q \cdot z > 0 \tag{3.13}$$

for all $z \neq 0$. Since $Qz = 0$ if and only if $z = 0$ our proof is completed.

□

It is important to mention that from Lemma 3.1 we can deduce that if A is a nonsingular square matrix then it is enough to minimize $\varphi(y)$ and the minimizer x^* will satisfy $x^* = A^{-1} \cdot y^*$.

In our next lemma we introduce a set of matrices $A \in R^{m \times n}$ such that for every factorization of a matrix in this set; $A = BQ$ where Q is a non-singular matrix, the matrix B will have at most $n^2 - n$ zeroes ($Z(B) \leq n^2 - n$). Next we show a practical method of factorizing a full ranked matrix which achieve at least $n^2 - n$ zeroes: in general we cannot expect more .

Lemma 3.2: Let $A \in R^{m \times n}$ where $m > n$ be a full rank matrix. Let $\bar{A} = [A, -I]$ be an m by $n+m$ matrix. If any set of m columns of $\bar{A}$ are linearly independent then for every factorization $A = BQ$ where $Q \in R^{n \times n}$ is a non-singular matrix and $B \in R^{m \times n}$, the matrix B will include , at least,

$n(m+1)-n^2$ nonzero elements. (that is, $Z(B) \le n^2-n$).

Proof: Consider the factorization $AQ^{-1} = B$ which can be written as n identical linear systems

$$A \cdot (Q^{-1})_{*j} - I \cdot B_{*j} = 0 \qquad j=1,\dots,n \tag{3.14}$$

The coefficients matrix $\tilde{A}=[A,-I]$ has rank m and any $m \times m$ submatrix of $\tilde{A}$ has full rank. Let us denote by x the vector $\begin{bmatrix} Q_{*j}^{-1} \\ B_{*j} \end{bmatrix}$ in R^{m+n}. First we claim that x has at least $m+1$ nonzero elements. Suppose x has less then $m+1$ nonzero elements then it has at least n zero elements. Suppose $x_{i_1}=x_{i_2}= \cdots =x_{i_n}=0$ and define $C \in R^{m \times m}$ to be a submatrix of $\tilde{A}$ with columns $\tilde{A}_{*j}$ where $j \ne i_k$ for all $1 \le k \le n$. According to the lemma's assumption, C is nonsingular and therefore the only solution to $C \cdot y = 0$ is $y=0$ which mean Q_{*j}^{-1} is zero. This contradicts our assumption that Q is nonsingular. Therefore the matrices Q and B together have at least $nm+n$ nonzero elements. If we assume that all the zeroes are in B, we still remain with $n(m+1)-n^2$ nonzero elements in B.

Comment: Any Vandermonde matrix satisfies the conditions of Lemma 3.2 therefore there are infinitely many examples of matrices for which one cannot expect to get more than $n^2 - n$ zeroes in B.

Next we introduce a practical method to factorize a full ranked matrix A with, at least, n^2-n zero elements in B.

The factorization

Let $A \in R^{m \times n}$ be a full rank matrix where $m > n$. Then we can write

$$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}. \tag{3.15}$$

Suppose A_1 is nonsingular $n \times n$ matrix. In this case we can take

$$B = \begin{bmatrix} I \\ A_2 \cdot A_1^{-1} \end{bmatrix} \qquad Q = [\, A_1 \,] \tag{3.16}$$

and there are $n^2 - n$ zeroes in B. However, this factorization is the worst case of section 1. In order to accomplish a better factorization, let as assume that $m > 2n$ in this case we can write the matrix as follow:

$$A = \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} \tag{3.17}$$

where $A_1 \in R^{n \times n}$ is a nonsingular matrix, $A_3 \in R^{n \times n}$ and $A_2 \in R^{(m-2n) \times n}$. Assume that $A_3 \cdot A_1^{-1}$ can be factorized into $L \cdot U$ where L and U are lower and upper triangular matrices respectively.

$$B = \begin{bmatrix} U^{-1} \\ A_2 \cdot A_1^{-1} \cdot U^{-1} \\ L \end{bmatrix} \qquad Q = U \cdot A_1 \tag{3.18}$$

will give us a factorization with $n^2 - n$ zeroes in B and its form will be closer to uniform distribution of the zero elements among the rows of the matrix.

It is interesting to observe cases in which the matrix A is not of full rank. We will show that in some cases it is possible to achieve more zeroes than the full rank case and in other cases, the opposite is true.

Lemma 3.3: Let $A \in R^{m \times n}$ where $rank(A) < n$. A sufficient condition for A to have a factorization $A = BQ$, where $Q \in R^{n \times n}$ is a nonsingular matrix and $B \in R^{m \times n}$ such that $Z(B) > n^2 - n$ is that

$$rank(A) + n < m + 1 \tag{3.19}$$

Proof: Suppose that $rank(A) = k$, $1 \leq k < n$. Without loss of generality we may assume that the first k columns of A are linearly independent and the last $(n-k)$ columns are linear combinations of the first k columns. Let us write $A = [A_1, A_2]$ where $A_1 \in R^{m \times k}$ and $A_2 \in R^{m \times (n-k)}$. There exists a matrix $E \in R^{k \times (n-k)}$ such that $A_2 = A_1 \cdot E$. The matrix A_1 can be factorized to $A_1 = B_1 \cdot Q_1$ according to (3.16) where $B_1 \in R^{m \times k}$ has $k^2 - k$ zero elements, and $Q_1 \in R^{k \times k}$ a nonsingular matrix. Let $B \in R^{m \times m}$ be the matrix with B_1 in its first K columns and zeroes in its last $(n-k)$ columns and let

$$Q = \begin{bmatrix} Q_1 & Q_1 \cdot E \\ 0 & I \end{bmatrix}. \tag{3.20}$$

Since Q_1 is nonsingular, Q is nonsingular and $A = BQ$. In this case B has at least $k^2 - k + m(n - k)$ zeroes. Recall that the number of zeroes in B in the full rank case is $n^2 - n$, it follows that $k^2 - k + m(n - k) > n^2 - n$ iff $k^2 - k(m + 1) + n(m + 1) - n^2 > 0$ iff $k^2 - n^2 > (m + 1)(k - n)$. Since $k < n$ the last inequality will hold iff $k + n < m + 1$. This inequality is the sufficient condition in (3.19).

Conclusions

We have seen in this paper a class of optimization problems for which the Hessian matrix can be written as $A^T \cdot D \cdot A$ where $A \in R^{m \times n}$ and $D \in R^{m \times m}$ a diagonal matrix. We showed that in several cases, the matrix A can be partially designed by the user in order to reduce the number of nonzero elements to a minimum. In previous sections we explored the pattern of a sparse matrix with a given number of nonzero elements. We showed that in order to minimize the computational complexity of $A^T \cdot D \cdot A$ we should divide the nonzero elements uniformly among the rows of A and if the nonzero elements are confined in certain rows then the computational complexity is maximized.

The difference between the evaluation of the product $A^T \cdot A$ by method of dense matrices and the upper bound for the worst case using sparse method is presented in (2.40). It can be seen that this difference depends linearly on the proportion of zero elements in the matrix which is $\frac{mn - N}{mn}$. Furthermore, the saving in using sparse method is, at least, $\frac{1}{2}(n+2)mn$ times this proportion. Since $\frac{1}{2}(n+2)mn$ and (2.38) are both close for large m and n, the saving is at least the number of operations for the dense case times the proportion of the zeroes elements.

Finally we demonstrated a practical method for factorizing a full ranked matrix $A \in R^{m \times n}$ into $B \cdot Q$ where B has at least $n^2 - n$ zero elements. Furthermore, we presented a class of matrices A for which you cannot expect to get more than $n^2 - n$ zero elements.

Unfortunately , this factorization is not optimal since the nonzero elements are not distributed uniformly among the rows and this question is still without an answer. Secondly, we proved that we can achieve at least $n^2 - n$ zero elements in B if A is full ranked or $rank(A) + n < m + 1$. We did not prove anything for matrices which are not full rank and do not satisfy (3.19) . The author conjecture is that the theorem may apply also for this case.

REFERENCES

1. GAY, D.M. **On Solving Robust and Generalized Linear Regression Problems.** in Optimizzazione Non-linear e Applicazioni, S. Incerti and G. Treccani, eds., Pitagora Editrice.

2. GEORGE, A. and LIU, J.W.; **Computer Solution of Large Sparse Positive Definite Systems,** Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1981

3. GONEN, A. and AVRIEL, M. ; **A Primal-Dual Newton-Type Algorithm for Geometric Programs with Equality Constraints** Journal of Optimization Theory and Applications, (to appear).

4. GUSTAVSON, F.G. **Two Fast Algorithms for Sparse Matrices: Multiplication and Permuted Transposition** ACM Transactions on Mathematical Software, Vol 4, No 3, pp. 250-269, 1978.

5. HOFFMAN, A.J. and McCORMICK, S.T. ; **A Fast Algorithm That Makes Matrices Optimally Sparse** Systems Optimization Laboratory, Department of Operations Research, Stanford University, Technical Report SOL 82-13 , September 1982.

DISTRIBUTION LIST

DEFENSE TECHNICAL INFORMATION CENTER (2)
CAMERON STATION
ALEXANDRIA, VIRGINIA 22214

CHIEF OF NAVAL RESEARCH (2)
ARLINGTON, VA 22217

LIBRARY, Code 0142 (2)
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943

RESEARCH ADMINISTRATION (1)
Code 012
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA 93943

PROFESSOR AMNON GONEN (18)
Code 53Gm
DEPARTMENT OF MATHEMATICS
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA 93943